

Fisher information and Hamilton's canonical equations

F. Pennini and A. Plastino

Instituto de Física La Plata (IFLP)

Universidad Nacional de La Plata (UNLP) and

Argentine National Research Council (CONICET)

C.C. 727, 1900 La Plata, Argentina

Abstract

We show that the mathematical form of the information measure of Fisher's I for a Gibbs' canonical probability distribution (the most important one in statistical mechanics) incorporates important features of the intrinsic structure of classical mechanics and *has a universal form* in terms of “forces” and “accelerations”, i.e., one that is valid for all Hamiltonian of the form $T + V$. If the system of differential equations associated to Hamilton's canonical equations of motion is linear, one can easily ascertain that the Fisher information per degree of freedom is proportional to the inverse temperature and to the number of these degrees. This equipartition of I is also seen to hold in a simple example involving a non-linear system of differential equations.

PACS numbers: 02.50.-r, 05.20.-y, 45.20.-d

INTRODUCTION

The last years have witnessed a great deal of activity revolving around physical applications of Fisher's information measure (FIM) (as a rather small sample, see for instance, [1, 2, 3, 4, 5]). Frieden and Soffer [1] have shown that Fisher's information measure provides one with a powerful variational principle, the extreme physical information one, that yields most of the canonical Lagrangians of theoretical physics [1, 2]. Additionally, I has been shown to provide an interesting characterization of the "arrow of time", alternative to the one associated with Boltzmann's entropy [6, 7]. Also to be mentioned is a recent approach to non-equilibrium thermodynamics, based upon Fisher's measure (a kind of "Fisher-MaxEnt"), that exhibits definite advantages over conventional text-book treatments [8, 9]. Thus, unravelling the multiple FIM facets and their links to physics should be of general interest to a vast audience. R. A. Fisher advanced, already in the twenties, a quite interesting information measure (for a detailed study see [1, 2]). Consider a $\theta - \mathbf{z}$ "scenario" in which we deal with a system specified by a physical parameter θ , while \mathbf{z} is a stochastic variable ($\mathbf{z} \in \mathbb{R}^M$) and $f_\theta(\mathbf{z})$ the probability density for \mathbf{z} (that depends also on θ). One makes a measurement of \mathbf{z} and has to best infer θ from this measurement, calling the resulting estimate $\tilde{\theta} = \tilde{\theta}(\mathbf{z})$. The question is how well θ can be determined. Estimation theory [10] states that the best possible estimator $\tilde{\theta}(\mathbf{z})$, after a very large number of \mathbf{z} -samples is examined, suffers a mean-square error e^2 from θ that obeys a relationship involving Fisher's I , namely, $I e^2 = 1$, where the Fisher information measure I is of the form

$$I(\theta) = \int d\mathbf{z} f_\theta(\mathbf{z}) \left\{ \frac{\partial \ln f_\theta(\mathbf{z})}{\partial \theta} \right\}^2. \quad (1)$$

This "best" estimator is the so-called *efficient* estimator. Any other estimator exhibits a larger mean-square error. The only caveat to the above result is that all estimators be unbiased, i.e., satisfy $\langle \tilde{\theta}(\mathbf{z}) \rangle = \theta$. Fisher's information measure has a lower bound: no matter what parameter of the system one chooses to measure, I has to be larger or equal than the inverse of the mean-square error associated with the concomitant experiment. This result, $I e^2 \geq 1$, is referred to as the Cramer-Rao bound [2]. A particular I -case is of great importance: that of translation families [2, 3], i.e., distribution functions (DF) whose *form* does not change under θ -displacements. These DF are shift-invariant (à la Mach, no absolute origin for θ), and for them Fisher's information measure adopts the somewhat

simpler appearance [2]

$$I(\text{shift-invariant}) = \int d\mathbf{z} f(\mathbf{z}) \left\{ \frac{\partial \ln f(\mathbf{z})}{\partial \mathbf{z}} \right\}^2. \quad (2)$$

Fisher’s measure is additive [2]. If one is dealing with phase-space, \mathbf{z} is a $M = 2N$ -dimensional vector, with N “coordinates” \mathbf{r} and N “momenta” \mathbf{p} . As a consequence, if one wishes to estimate phase-space “location”, one confronts an $I(\mathbf{r} + \mathbf{p}) \equiv I(\mathbf{z})$ -measure and [11]

$$I(\mathbf{z}) = I(\mathbf{r}) + I(\mathbf{p}). \quad (3)$$

Here we clearly encounter a shift-invariance situation.

HAMILTONIAN SYSTEMS

Assume we wish to describe a classical system of N identical particles of mass m whose Hamiltonian is of the form (in self-explanatory notation)

$$\mathcal{H} = \mathcal{T} + \mathcal{V}. \quad (4)$$

A special important instance is that of

$$\mathcal{H}(\mathbf{r}, \mathbf{p}) = \sum_{i=1}^N \frac{\mathbf{p}_i^2}{2m} + \sum_{i=1}^N V(r_i) \quad (5)$$

where $(\mathbf{r}_i, \mathbf{p}_i)$ are the coordinates and momenta of the i -th particle and $V(\mathbf{r}_i)$ is a central single-particle potential. Our considerations are *not limited* to Hamiltonians (5), though.

We presuppose that the system is in equilibrium at temperature T , so that, in the canonical ensemble, the probability density reads

$$\rho(\mathbf{r}, \mathbf{p}) = \frac{e^{-\beta \mathcal{H}(\mathbf{r}, \mathbf{p})}}{Z} \quad (6)$$

with k Boltzmann’s constant and $\mathbf{r} - \mathbf{p}$ denoting two $3N$ -dimensional vectors. If h denotes an elementary cell in phase-space, we write, with some abuse of notation [12], $d\tau \equiv d^{3N}\mathbf{r} d^{3N}\mathbf{p} / (N! h^{3N})$ for the pertinent integration-measure while the partition function reads [12]

$$Z = \int \frac{d^{3N}\mathbf{r} d^{3N}\mathbf{p}}{N! h^{3N}} e^{-\beta \mathcal{H}(\mathbf{r}, \mathbf{p})}. \quad (7)$$

Remember now Hamilton's celebrated canonical equations

$$\frac{\partial \mathcal{H}(\mathbf{r}, \mathbf{p})}{\partial \mathbf{p}} = \dot{\mathbf{r}}; \quad \frac{\partial \mathcal{H}(\mathbf{r}, \mathbf{p})}{\partial \mathbf{r}} = -\dot{\mathbf{p}}, \quad (8)$$

so that, obviously,

$$\begin{aligned} -kT \frac{\partial \ln \rho(\mathbf{r}, \mathbf{p})}{\partial \mathbf{p}} &= \dot{\mathbf{r}} \\ -kT \frac{\partial \ln \rho(\mathbf{r}, \mathbf{p})}{\partial \mathbf{r}} &= -\dot{\mathbf{p}}. \end{aligned} \quad (9)$$

Eqs. (9) are just the above referred to canonical equations of Hamilton in a different guise. We recall them here because the expressions in the left-hand-sides of (9) enter the definition of Fisher's measure, if one expresses I_τ as a functional of the probability density ρ . Indeed, this FIM has the form (2), so that, in view of (3), it acquires the aspect

$$I_\tau = \int \frac{d^{3N} \mathbf{r} d^{3N} \mathbf{p}}{N! h^{3N}} \rho(\mathbf{r}, \mathbf{p}) \mathcal{A}(\mathbf{r}, \mathbf{p}) \quad (10)$$

with

$$\mathcal{A} = a \left[\frac{\partial \ln \rho(\mathbf{r}, \mathbf{p})}{\partial \mathbf{p}} \right]^2 + b \left[\frac{\partial \ln \rho(\mathbf{r}, \mathbf{p})}{\partial \mathbf{r}} \right]^2 \quad (11)$$

where, for the sake of dimensional balance [11], one needs two coefficients a and b that depend on the specific Hamiltonian of the system under scrutiny (see the following Section).

Notice that

$$\frac{\partial \ln \rho(\mathbf{r}, \mathbf{p})}{\partial \mathbf{p}} = -\beta \frac{\partial \mathcal{H}(\mathbf{r}, \mathbf{p})}{\partial \mathbf{p}} \quad (12)$$

and

$$\frac{\partial \ln \rho(\mathbf{r}, \mathbf{p})}{\partial \mathbf{r}} = -\beta \frac{\partial \mathcal{H}(\mathbf{r}, \mathbf{p})}{\partial \mathbf{r}} \quad (13)$$

so that we can rewrite the Fisher information as

$$(kT)^2 I_\tau = a \left\langle \left(\frac{\partial \mathcal{H}(\mathbf{r}, \mathbf{p})}{\partial \mathbf{p}} \right)^2 \right\rangle + b \left\langle \left(\frac{\partial \mathcal{H}(\mathbf{r}, \mathbf{p})}{\partial \mathbf{r}} \right)^2 \right\rangle. \quad (14)$$

In view of Eqs. (9), Fisher's measure now becomes

$$I_\tau = \beta^2 \left\{ a \langle \dot{\mathbf{r}}^2 \rangle + b \langle \dot{\mathbf{p}}^2 \rangle \right\}. \quad (15)$$

We see that it incorporates the symplectic structure of classical mechanics. *Eqs. (14-15) are our main original results* here: they give the FIM-form for any Hamiltonian of the form (4),

a universal “quadratic” form in the average values of “square-accelerations” and “square-forces”. The first term in the right-hand-side of (14) is in fact the same for all systems for which \mathcal{T} is the conventional kinetic energy. Note that cyclic canonical coordinates do not contribute to I_τ .

SOME SPECIALLY IMPORTANT EXAMPLES: GENERAL QUADRATIC HAMILTONIANS

In many important cases [12] one faces Hamiltonians of the type

$$\mathcal{H}_L = \sum_{i=1}^N \left(A_i P_i^2 + B_i Q_i^2 \right), \quad (16)$$

with $P_i - Q_i$ generalized momenta and coordinates, respectively, that lead to $2N$ differential equations, the so-called linear Hamiltonian problem [13]. According to the preceding section we have now

$$(kT)^2 I_\tau = \sum_{i=1}^N \left[\tilde{A}_i \left\langle \left(\frac{\partial \mathcal{H}_L}{\partial P_i} \right)^2 \right\rangle + \tilde{B}_i \left\langle \left(\frac{\partial \mathcal{H}_L}{\partial Q_i} \right)^2 \right\rangle \right], \quad (17)$$

it being then obvious that, in order to achieve a proper dimensional balance, we must have $\tilde{A}_i = 1/(\beta_0 A_i)$ and $\tilde{B}_i = 1/(\beta_0 B_i)$ for all i , where β_0 is a fixed, arbitrary reference inverse-temperature. Further reflection allows one to see that

$$(kT)^2 I_\tau = \langle \mathcal{H}_L \rangle / \beta_0. \quad (18)$$

We know that the equipartition theorem holds for $\langle \mathcal{H}_L \rangle$ [12], so that the mean value of the energy equals $2NkT$, which entails that it also holds for the Fisher information

$$I_\tau = 2N (\beta / \beta_0). \quad (19)$$

As expected, the information steadily diminishes as the temperature grows. The divergence at zero temperature is of no importance, since classical mechanics is not valid at low enough temperatures.

Free particle

In the case of the N identical free particles, the Hamiltonian is $\mathcal{H}_i = \mathbf{p}_i^2/2m$ with $i = 1 \dots N$. We make here the obvious choice $a = m/\beta_0$ (mass). Since $\langle \dot{\mathbf{r}}^2 \rangle = 2\langle \mathcal{H} \rangle/m$ we have

$$I_\tau = 2\beta^2 \langle \mathcal{H}/\beta_0 \rangle, \quad (20)$$

so that, using the well-known result (equipartition) [12] $\langle \mathcal{H} \rangle = 3N/2\beta$ one obtains

$$I_\tau = 3N (\beta/\beta_0). \quad (21)$$

N harmonic oscillators

We consider N identical harmonic oscillators whose Hamiltonian is $\mathcal{H}_i = \mathbf{p}_i^2/2m + m\omega \mathbf{r}_i^2/2$ with $i = 1 \dots N$. It is easy to calculate the classical mean values in (15), which have the form $\langle \dot{\mathbf{r}}^2 \rangle = 3N/m\beta$ and $\langle \dot{\mathbf{p}}^2 \rangle = 3Nm\omega^2/\beta$. In this case the proper choice is $a = m/\beta_0$ and $b = 1/m\beta_0\omega^2$, so that the information (15) adopts the appearance

$$I_\tau = 6N (\beta/\beta_0). \quad (22)$$

In these two examples we appreciate again the fact that a Fisher–equipartition emerges, kT/β_0 per degree of freedom.

PARAMAGNETIC SYSTEM

Consider N magnetic dipoles, each of magnetic moment μ , in the presence of an external magnetic field \mathbf{H} of intensity H . These N distinguishable (localized), identical, mutually non-interacting, and freely orientable dipoles give rise to a dipole-potential energy [12]

$$\mathcal{H} = - \sum_{i=1}^N \vec{\mu}_i \cdot \mathbf{H} = -\mu H \sum_{i=1}^N \cos \theta_i, \quad (23)$$

where θ_i gives the dipole orientation with respect to the field-direction. Since there is no interaction between our N particles, and both I and the entropy S are additive quantities, it is permissible to focus attention on just one generic particle (whose canonical conjugate variables we call (θ, p_θ)) and, at the end of the calculation, multiply by N . Hamilton’s canonical equations yield then the non-linear equation

$$\dot{p}_\theta = \mu H \sin \theta. \quad (24)$$

Notice that p_θ does not appear in \mathcal{H} , i.e., $\partial\mathcal{H}/\partial p_\theta = \dot{\theta} = 0$, which entails, of course, $\langle\dot{\theta}\rangle = 0$. Thus, in choosing the constants entering the definition of I_τ (see the preceding Section), we need only to care about b , with $b = 1/(\mu\beta_0 H)$. The associated Fisher information is then (Cf. Eq. (15))

$$I_\tau = \frac{\beta^2 N}{\mu H} \langle \sin^2 \theta \rangle. \quad (25)$$

We can easily compute the above mean value using $(\sin\theta \, d\theta \, d\varphi)$ for the elemental solid angle, so that

$$\langle \sin^2 \theta \rangle = \frac{1}{Z} \int_0^{2\pi} \int_0^\pi e^{\beta\mu H} \sin^3 \theta \, d\theta \, d\varphi \quad (26)$$

where the partition function per particle is of the form [12]

$$Z = 4\pi \frac{\sinh(x)}{x}, \quad (27)$$

with $x = \mu H/kT$. Evaluating the integral (26) explicitly we obtain

$$I_\tau = 2 N (\beta/\beta_0) L(x), \quad (28)$$

an *original* result, where $L(x)$ is the well known *Langevin function* [12]

$$L(x) = \coth x - \frac{1}{x}. \quad (29)$$

Since $L(x)$ vanishes at the origin, so does I_τ for infinite temperature, a result one should expect on more general grounds [11]. Again, equipartition of Fisher information ensues. Since the differential equations are not linear in the conjugate (canonical) variables, the information per degree of freedom does not equal that of (19).

CONCLUSIONS

We summarize now our results. In this communication we have shown that

- the mathematical form of the information measure of Fisher's I for a Gibbs' canonical probability distribution incorporates important features of the intrinsic structure of classical mechanics
- it has a universal form in terms of “forces” and “accelerations”, i.e., one that is valid for any Hamiltonian of the form $T + V$

- if the system of differential equations associated to Hamilton's canonical equations of motion is linear, the amount of Fisher information per degree of freedom is proportional to the inverse temperature and to the number of these degrees (equipartition of information!)
- equipartition of I has been seen to hold also for paramagnetic systems, for which a non-linear system of differential equations is involved.

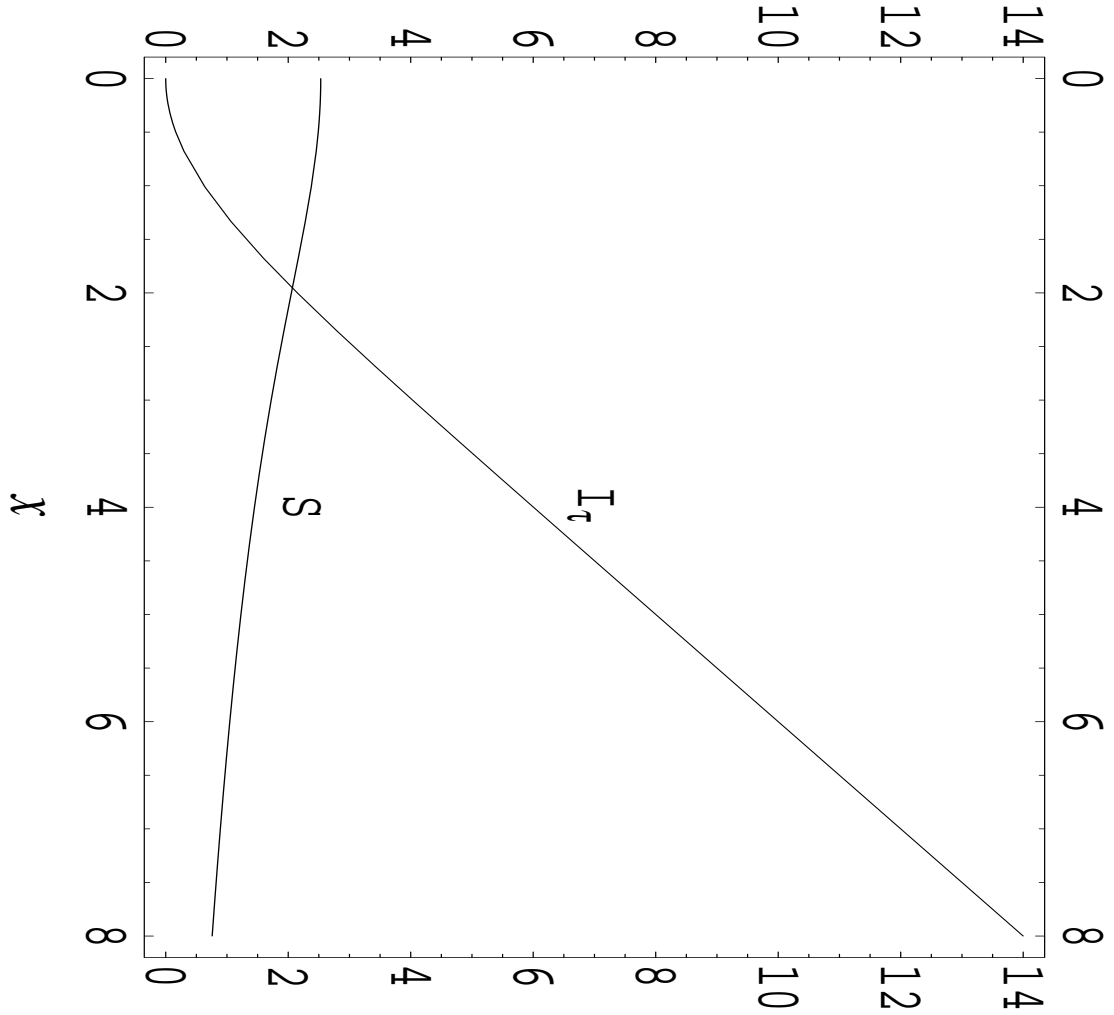


FIG. 1: We plot i) the entropy S and ii) the Fisher information I_τ per particle vs. $x = \mu H/kT$ for N magnetic dipoles, each of magnetic moment μ , in the presence of an external magnetic field \mathbf{H} of intensity H .

-
- [1] B.R. Frieden and B.H. Soffer, *Phys. Rev. E* **52**, 2274 (1995).
 - [2] B.R. Frieden, *Physics from Fisher information* (Cambridge University Press, Cambridge, England, 1998).
 - [3] F. Pennini, A.R. Plastino and A. Plastino, *Physica A* **258**, 446 (1998).
 - [4] B.R. Frieden, A. Plastino, A.R. Plastino and H. Soffer, *Physical Review E* **60**, 48 (1999).
 - [5] F. Pennini, A. Plastino, A.R. Plastino and M. Casas, *Physics Letters A* **302**, 156 (2002).
 - [6] A.R. Plastino and A. Plastino, *Phys. Rev. E* **54**, 4423 (1996).
 - [7] A. Plastino, A.R. Plastino, and H.G. Miller, *Phys. Lett. A* **235**, 129 (1997).
 - [8] R. Frieden, A. Plastino, A. R. Plastino, and B. H. Soffer, *Phys. Rev. E* **60**, 48 (1999); **66**, 046128 (2002); *Phys. Lett. A* **304**, 73(2002).
 - [9] S. Flego, R. Frieden, A. Plastino, A. R. Plastino, and B. H. Soffer, *The Physical Review E* **68**, 016105 (2003).
 - [10] H. Cramer, *Mathematical methods of statistics*, (Princeton University Press, Princeton, NJ, 1946).
 - [11] F. Pennini and A. Plastino, *Phys. Rev. E* (2004, in Press); preprint (2003) [cond-mat/0312680].
 - [12] R.K. Pathria, *Statistical Mechanics* (Pergamon Press, Exeter, 1993).
 - [13] A.J. Lichtenberg and M.A. Lieberman, *Regular and chaotic dynamics*, Springer-Verlag (Berlin), (1991).